EXHIBIT G

TO DECLARATION OF S. MERRILL WEISS IN SUPPORT OF PLAINTIFF ACACIA MEDIA TECHNOLOGIES CORPORATION'S MEMORANDUM OF POINTS AND AUTHORITIES IN OPPOSITION TO ROUND 3 DEFENDANTS' MOTION FOR SUMMARY JUDGMENT OF INVALIDITY UNDER 35 U.S.C. § 112 OF THE '992, '863, AND '702 PATENTS; AND SATELLITE DEFENDANTS' MOTION FOR SUMMARY JUDGMENT OF INVALIDITY OF THE '992, '863, AND '720 PATENTS

Hovsepian v. Apple, Inc. Doc. 325 Att. 6

Dockets.Justia.com

IEEE TRANSACTIONS ON

COMPUTERS

JANUARY 1974 VOLUME C-23 NUMBER 1

A PUBLICATION OF THE IEEE COMPUTER SOCIETY

UNIVERSITY OF CALIFORNIA
LOS ANGELES
SEP 8 1975
ENGINEERING &
MATHEMATICAL
SCIENCES LIBRARY

PAPERS

Algorithms

Fast Computational Algorithms for Bit Reversal .... *R. J. Polge, B. K. Bhagavan, and J. M. Carswell* 1

Arithmetic Systems

Floating-Point Arithmetic in the Symmetric Residue System .... *E. Kinoshita, H. Kosako, and Y. Kojima* 9

Asynchronous Machines

Towards a Theory of Universal Speed-Independent Modules .... *R. M. Keller* 21

Combinational Networks

On Two-Level EXCLUSIVE-OR Majority Networks .... *W. E. Robbins and L. D. Rudolph* 34

Fault Tolerant Machines

Design of Two-Level Fault-Tolerant Networks .... *D. K. Pradhan and S. M. Reddy* 41

Multiple Fault Detection and Location in Fan-Out-Free Combinational Circuits .... *G. Fantauzzi and A. Marsella* 48

On Modifying Logic Networks to Improve Their Diagnosability .... *J. P. Hayes* 56

Logic Design

Optimal One-Bit Full Adders with Different Types of Gates .... *T. K. Liu, K. Hohulin, L. E. Shiau, and S. Muroga* 63

Pattern Recognition

On a New Class of Bayes Risk in Multihypothesis Pattern Recognition .... *P. A. Devijver* 70

Sequential Machines

State Minimization of Incompletely Specified Sequential Machines .... *N. N. Biswas* 80

CORRESPONDENCE

A State Assignment Technique for Sequential Machines Using *J-K* Flip-Flops .... *G. L. Tumbush and J. E. Brandeberry* 85

Characterization of Connection Assignment of Diagnosable Systems .... *S. L. Hakimi and A. T. Amin* 86

A New Algorithm for Computing Correlations .... *W. B. Kendall* 88

Discrete Cosine Transform .... *N. Ahmed, T. Natarajan, and K. R. Rao* 90

Minimization of Linear Sequential Machines .... *C.-T. Chen* 93

Sequential Algorithm for the Determination of Maximum Compatibles .... *K. E. Stoffers* 95

Easily Testable Cellular Realizations for the (Exactly *P*)-out-of *n* and (*p* or More)-out-of *n* Logic Functions .... *S. M. Reddy and J. R. Wilson* 98

A Computational Algorithm for Solving a System of Coupled Algebraic Matrix Riccati Equations .... *A. I. A. Salama and V. Gourishankar* 100

Quantization Complexity and Independent Measurements .... *B. Chandrasekaran and A. K. Jain* 102

A Novel Implementation Method for Addition and Subtraction in Residue Number Systems .... *D. K. Banerji* 106

BOOK REVIEWS

B74-1 Complexity of Computer Computations—R. E. Miller and J. W. Thatcher, Ed. .... *Reviewed by J. R. Rice* 109

B74-2 Man and the Computer—J. G. Kenemy .... *Reviewed by T. H. Bonn* 110

B74-3 Design Automation of Digital Systems—Volume 1 Theory Techniques—M. A. Bruer, Ed. .... *Reviewed by H. Y. Chang* 110

B74-4 Minicomputers in Data Processing and Simulation—B. Soucek .... *Reviewed by J. D. Schoeffler* 111

B74-5 Computer-Oriented Approaches to Pattern Recognition—W. S. Meisel .... *Reviewed by R. C. Gillepie* 112

If only the L values R_0 through R_{L-1} are required, then it can be seen by comparing (15) and (11) that the new algorithm will require fewer multiplications than the FFT method if

$$L < 11.2\,(1 + \log_2 N) \tag{16a}$$

and will require fewer additions if

$$L < 5.6\,(1 + \log_2 N). \tag{16b}$$

Therefore, we conclude that the new algorithm will generally be more efficient than the FFT method if

$$N \leq 128 \tag{17a}$$

or

$$L \leq 10(1 + \log_2 N). \tag{17b}$$

Conclusion

A new algorithm for computing the correlation of a block of sampled data has been presented. It is a direct method which trades an increased number of additions for a decreased number of multiplications. For applications where the "cost" (e.g., the time) of a multiplication is greater than that of an addition, the new algorithm is always more computationally efficient than direct evaluation of the correlation, and it is generally more efficient than FFT methods for processing 128 or fewer data points, or for calculating only the first L "lags" for $L \leq 10 \log_2 2N$.

References

[1] S. Winograd, "A new algorithm for inner product," *IEEE Trans. Comput.* (Corresp.), vol. C-17, pp. 693–694, July 1968.

Discrete Cosine Transform

N. AHMED, T. NATARAJAN, AND K. R. RAO

Abstract—**A discrete cosine transform (DCT) is defined and an algorithm to compute it using the fast Fourier transform is developed. It is shown that the discrete cosine transform can be used in the area of digital processing for the purposes of pattern recognition and Wiener filtering. Its performance is compared with that of a class of orthogonal transforms and is found to compare closely to that of the Karhunen–Loève transform, which is known to be optimal. The performances of the Karhunen–Loève and discrete cosine transforms are also found to compare closely with respect to the rate-distortion criterion.**

Index Terms—**Discrete cosine transform, discrete Fourier transform, feature selection, Haar transform, Karhunen–Loève transform, rate distortion, Walsh-Hadamard transform, Wiener vector and scalar filtering.**

Introduction

In recent years there has been an increasing interest with respect to using a class of orthogonal transforms in the general area of digital signal processing. This correspondence addresses itself towards two problems associated with image processing, namely, pattern recognition [1] and Wiener filtering [2].

In pattern recognition, orthogonal transforms enable a noninvertible transformation from the pattern space to a reduced dimensionality feature space. This allows a classification scheme to be implemented with substantially less features, with only a small increase in classification error.

Manuscript received January 29, 1973; revised June 7, 1973.
N. Ahmed is with the Departments of Electrical Engineering and Computer Science, Kansas State University, Manhattan, Kans.
T. Natarajan is with the Department of Electrical Engineering, Kansas State University, Manhattan, Kans.
K. R. Rao is with the Department of Electrical Engineering, University of Texas at Arlington, Arlington, Tex. 76010.

In discrete Wiener filtering applications, the filter is represented by an $(M \times M)$ matrix G. The estimate $\hat{X}$ of data vector X is given by GZ, where $Z = X + N$ and N is the noise vector. This implies that approximately $2M^2$ arithmetic operations are required to compute $\hat{X}$. Use of orthogonal transforms yields a G in which a substantial number of elements are relatively small in magnitude, and hence can be set equal to zero. Thus a significant reduction in computation load is realized at the expense of a small increase in the mean-square estimation error.

The Walsh-Hadamard transform (WHT), discrete Fourier transform (DFT), the Haar transform (HT), and the slant transform (ST), have been considered for various applications [1], [2], [4]–[9] since these are orthogonal transforms that can be computed using fast algorithms. The performance of these transforms is generally compared with that of the Karhunen–Loève transform (KLT) which is known to be optimal with respect to the following performance measures: variance distribution [1], estimation using the mean-square error criterion [2], [4], and the rate-distortion function [5]. Although the KLT is optimal, there is no general algorithm that enables its fast computation [1].

In this correspondence, a discrete cosine transform (DCT) is introduced along with an algorithm that enables its fast computation. It is shown that the performance of the DCT compares more closely to that of the KLT relative to the performances of the DFT, WHT, and HT.

Discrete Cosine Transform

The DCT of a data sequence $X(m)$, $m = 0, 1, \cdots, (M-1)$ is defined as

$$G_x(0) = \frac{\sqrt{2}}{M} \sum_{m=0}^{M-1} X(m)$$

$$G_x(k) = \frac{2}{M} \sum_{m=0}^{M-1} X(m) \cos \frac{(2m+1)k\pi}{2M}, \qquad k = 1, 2, \cdots, (M-1) \tag{1}$$

where $G_x(k)$ is the kth DCT coefficient. It is worthwhile noting that the set of basis vectors $\{1/\sqrt{2}, \cos((2m+1)k\pi)/(2M)\}$ is actually a class of discrete Chebyshev polynomials. This can be seen by recalling that Chebyshev polynomials can be defined as [3]

$$\hat{T}_0(\xi_p) = \frac{1}{\sqrt{2}}$$

$$\hat{T}_k(\xi_p) = \cos(k \cos^{-1} \xi_p), \qquad k, p = 1, 2, \cdots, M \tag{2}$$

where $\hat{T}_k(\xi_p)$ is the kth Chebyshev polynomial.

Now, in (2), ξ_p is chosen to be the pth zero of $\hat{T}_M(\xi)$, which is given by [3]

$$\xi_p = \cos \frac{(2p-1)\pi}{2M}, \qquad p = 1, 2, \cdots, M. \tag{3}$$

Substituting (3) in (2), one obtains the set of Chebyshev polynomials

$$\hat{T}_0(p) = \frac{1}{\sqrt{2}}$$

$$\hat{T}_k(p) = \cos \frac{(2p-1)k\pi}{2M}, \qquad k, p = 1, 2, \cdots, M. \tag{4}$$

From (4) it follows that the $\hat{T}_k(p)$ can equivalently be defined as

$$T_0(m) = \frac{1}{\sqrt{2}}$$

$$T_k(m) = \cos \frac{(2m+1)k\pi}{2M}, \qquad k = 1, 2, \cdots, (M-1),$$

$$m = 0, 1, \cdots, M-1. \tag{5}$$

Comparing (5) with (1) we conclude that the basis member cos ((2m +

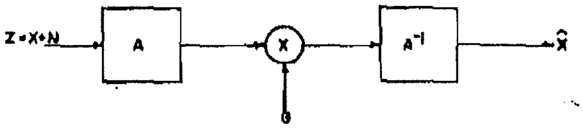

Fig. 3. Wiener filtering model.

TABLE I
MEAN-SQUARE ERROR PEFORMANCE OF VARIOUS TRANSFORMS FOR SCALAR WIENER FILTERING; $\rho = 0.9$

Transform \ M	2	4	8	16	32	64
Karhunen-Loève	0.3730	0.2915	0.2533	0.2356	0.2268	0.2224
Discrete cosine	0.3730	0.2920	0.2546	0.2374	0.2282	0.2232
Discrete Fourier	0.3730	0.2964	0.2706	0.2592	0.2441	0.2320
Walsh-Hadamard	0.3730	0.2942	0.2649	0.2582	0.2582	0.2559
Haar	0.3730	0.2942	0.2650	0.2589	0.2582	0.2581

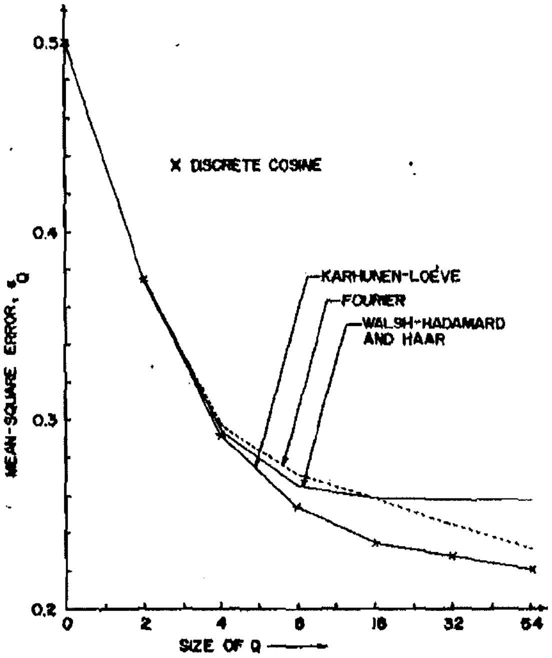

Fig. 4. Mean-square error performance of various transforms for scalar Wiener filtering; $\rho = 0.9$.

vector X and a noise vector N. X is considered to belong to a random process whose covariance matrix is given by ψ which is defined in (9). The Wiener filter G is in the form of an $(M \times M)$ matrix. A and A^{-1} represent an orthonormal transform and its inverse, respectively, while $\hat{X}$ denotes the estimate of X, using the mean-square error criterion.

We restrict our attention to the case when G is constrained to be a diagonal matrix Q. This class of Wiener filters is referred to as *scalar* filters while the more general class (denoted by G) is referred to as vector filters. *The additive noise (see Fig. 3) is considered to be white,* zero mean, and uncorrelated with the data. If the mean-square estimation error due to scalar filtering is denoted by e_Q, then e_Q can be expressed as [4]

$$e_Q = 1 - \frac{1}{M}\sum_{s=1}^{M} \frac{\Psi_x^2(s,s)}{\Psi_x(s,s)+\Psi_n(s,s)} \tag{13}$$

where Ψ_x and Ψ_n denote the transform domain covariance matrices of the data and noise, respectively. Table I lists the values of e_Q for different values of M for the case $\rho = 0.9$ and a signal-to-noise ratio of unity. From Table I it is evident that the DCT comes closest to the KLT which is optimal. This information is presented in terms of a set of performance curves in Fig. 4.

ADDITIONAL CONSIDERATIONS

In conclusion, we compare the performance of the DCT with KLT, DFT, and the identity transforms, using the rate-distortion criterion [5]. This performance criterion provides a measure of the information rate R that can be achieved while still maintaining a fixed distortion D, for encoding purposes. Considering Gaussian sources along with the mean-square error criterion, the rate-distortion performance measure is given by [5]

$$R(A,D) = \frac{1}{2M}\sum_{j=1}^{M} \max\left\{0, \ln\left(\frac{\sigma_j}{\theta}\right)\right\} \tag{14a}$$

Exhibit 9 Page 172

$1)k\Pi)/(2M)$ is the kth Chebyshev polynomial $T_k(\xi)$ evaluated at the mth zero of $T_M(\xi)$.

Again, the inverse cosine discrete transform (ICDT) is defined as

$$X(m) = \frac{1}{\sqrt{2}} G_x(0) + \sum_{k=1}^{M-1} G_x(k) \cos \frac{(2m+1)k\Pi}{2M}, \quad m = 0, 1, \cdots, (M-1). \quad (6)$$

We note that applying the orthogonal property [3]

$$\sum_{m=0}^{M-1} T_k(m)\, T_l(m) = \begin{cases} M/2, & k = l = 0 \\ M/2, & k = l \neq 0 \\ 0, & k \neq l \end{cases} \quad (7)$$

to (6) yields the DCT in (1).

If (6) is written in matrix form and Λ is the $(M \times M)$ matrix that denotes the cosine transformation, then the orthogonal property can be expressed as

$$\Lambda^T \Lambda = \frac{M}{2} [I] \quad (8)$$

where Λ^T is the transpose of Λ and $[I]$ is the $(M \times M)$ identity matrix.

Motivation

The motivation for defining the DCT is that it can be demonstrated that its basis set provides a good approximation to the eigenvectors of the class of Toeplitz matrices defined as

$$\psi = \begin{bmatrix} 1 & \rho & \rho^2 & & \rho^{M-1} \\ \rho & 1 & \rho & \cdots & \rho^{M-2} \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \rho^{M-1} & \rho^{M-2} & & \cdots & 1 \end{bmatrix}, \quad 0 < \rho < 1. \quad (9)$$

For the purposes of illustration, the eigenvectors of ψ for $M = 8$ and $\rho = 0.9$ are plotted (see Fig. 1) against

$$\left\{ \frac{1}{\sqrt{2}}, \cos \frac{(2m+1)k\Pi}{16}, \quad k = 1, 2, \cdots, 7, m = 0, 1, \cdots, 7 \right\} \quad (10)$$

which constitute the basis set for the DCT. The close resemblance (aside from the 180° phase shift) between the eigenvectors and the set defined in (10) is apparent.

Algorithms

It can be shown that (1) can be expressed as

$$G_x(0) = \frac{\sqrt{2}}{M} \sum_{m=0}^{M-1} X(m)$$

$$G_x(k) = \frac{2}{M} \operatorname{Re} \left\{ e^{(-ik\Pi)/(2M)} \sum_{m=0}^{2M-1} X(m)\, W^{km} \right\}, \quad k = 1, 2, \cdots, (M-1) \quad (11)$$

where

$$W = e^{-i2\Pi/2M}, \quad i = \sqrt{-1},$$

$$X(m) = 0, \quad m = M, (M+1), \cdots, (2M-1)$$

and Re $\{\cdot\}$ implies the real part of the term enclosed. From (11) it follows that all the M DCT coefficients can be computed using a $2M$-point fast Fourier transform (FFT). Since (1) and (6) are of the same form, FFT can also be used to compute the IDCT. Similarly, if a discrete sine transform were defined, then the Re $\{\cdot\}$ in (11) would be replaced by Im $\{\cdot\}$, which denotes the imaginary part of the term enclosed.

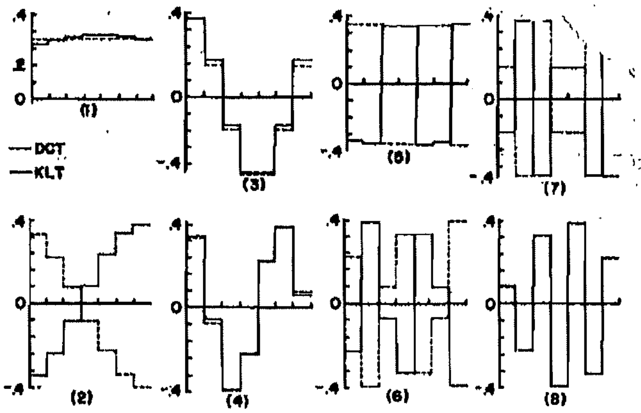

Fig. 1. Eigenvectors of (8 × 8) Toeplitz matrix (ρ = 0.9) and basis vectors of the DCT.

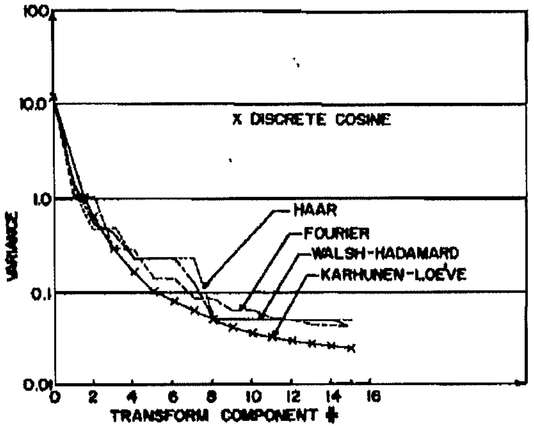

Fig. 2. Transform domain variance; M = 16, ρ = 0.95.

Computational Results

In image processing applications, ψ in (9) provides a useful model for the data covariance matrix corresponding to the rows and columns of an image matrix [6], [7]. The covariance matrix in the transform domain is denoted by Ψ and is given by

$$\Psi = \Lambda \psi \Lambda^{*T} \quad (12)$$

where Λ is the matrix representation of an orthogonal transformation and Λ^* is the complex conjugate of Λ. From (12) it follows that Ψ can be computed as a two-dimensional transform of ψ.

Feature Selection

A criterion for eliminating features (i.e., components of a transform vector), which are least useful for classification purposes, was developed by Andrews [1]. It states that features whose variances (i.e., main diagonal elements of Ψ) are relatively large should be retained. (Fig. 2 should be retained.) Fig. 2 shows the various variances ranked in decreasing order of magnitude. From the information in Fig. 2 it is apparent that relative to the set of orthogonal transforms shown, the DCT compares most closely to the KLT.

Wiener Filtering

The role of orthogonal transforms played in filtering applications is illustrated in Fig. 3 [2]. Z is an $(M \times 1)$ vector which is the sum of a

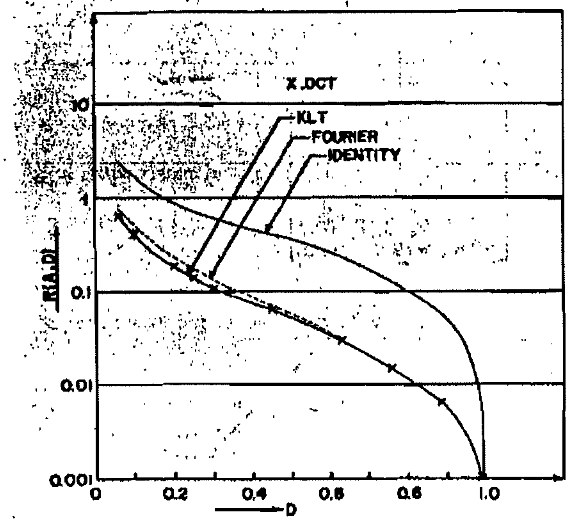

Fig. 5. Rate versus distortion for $M = 16$ and $\rho = 0.9$.

$$D = \frac{1}{M} \sum_{j=1}^{M} \min(\theta, \sigma_j) \tag{14b}$$

where A denotes the orthogonal transformation and the σ_j are the main diagonal terms of the transform domain covariance matrix Ψ in (12).

The rate-distortion pertaining to $M = 16$ and $\rho = 0.9$ is shown in Fig. 5, from which it is evident that the KLT and DCT compare more closely than the KLT and DFT.

Summary

It has been shown that the DCT can be used in the area of image processing for the purposes of feature selection in pattern recognition; and scalar-type Wiener filtering. Its performance compares closely with that of the KLT, which is considered to be optimal. The performances of the KLT and DCT are also found to compare closely, with respect to the rate-distortion criterion.

Acknowledgment

The authors wish to thank Dr. H. C. Andrews, University of Southern California, for his valuable suggestions pertaining to the rate-distortion computation.

References

[1] H. C. Andrews, "Multidimensional rotations in feature selection," *IEEE Trans. Comput.*, vol. C-20, pp. 1045–1051, Sept. 1971.
[2] W. K. Pratt, "Generalized Wiener filtering computation techniques," *IEEE Trans. Comput.*, vol. C-21, pp. 636–641, July 1972.
[3] C. T. Fike, *Computer Evaluation of Mathematical Functions*. Englewood Cliffs, N. J.: Prentice-Hall, 1968, ch. 6 and 7.
[4] J. Pearl, "Walsh processing of random signals," *IEEE Trans. Electromag. Compat.*, vol. EMC-13, pp. 137–141, Aug. 1971.
[5] J. Pearl, H. C. Andrews, and W. K. Pratt, "Performance measures for transform data coding," *IEEE Trans. Commun. Technol.*, vol. COM-20, pp. 411–415, June 1972.
[6] W. K. Pratt, "Walsh functions in image processing and two-dimensional filtering," in *1972 Proc. Symp. Applications of Walsh Functions*, AD-744 650, pp. 14–22.
[7] W. K. Pratt, L. R. Welch, and W. Chen, "Slant transforms in image coding," in *1972 Proc. Symp. Applications of Walsh Functions*, pp. 229–234.
[8] W. H. Chen and W. K. Pratt, "Color image coding with the slant transform," in *1973 Proc. Symp. Applications of Walsh Functions*, AD-763000, pp. 155–161.
[9] W. K. Pratt, "Spatial transform coding of color images," *IEEE Trans. Commun. Technol.*, vol. COM-19, pp. 980–992, Dec. 1971.

Minimization of Linear Sequential Machines

CHI-TSONG CHEN

Abstract—An algorithm is presented to minimize linear sequential machines to reduced form.

Index Terms—Algorithm, linear machine, minimization, reduced machine, reduction, sequential machine.

I. Introduction

The minimization of a linear sequential machine to a reduced form is an important topic and is discussed in many texts [1]–[4]. The minimization procedure presented in [1]–[4] is as follows. Let $\{A, B, C, D\}$ be an n-dimensional linear machine over $GF(p)$, and let r, with $r < n$, be the rank of the diagnostic matrix $K \triangleq [C'A'C' \cdots (A')^{n-1}C']'$, where the prime stands for the transpose. Define an $r \times n$ matrix T consisting of the first r linearly independent rows of K, and an $n \times r$ matrix R denoting the right inverse of T so that $TR = I$. Then the linear machine $\{TAR, TB, CR, D\}$ is a reduced form of $\{A, B, C, D\}$. In this correspondence, an algorithm will be introduced to find a special set of r linearly independent rows in K. A reduced machine can then be read out from this algorithm without the need of inverting any matrix. Furthermore, the reduced machine will be in a canonical form.

II. Algorithm

Let C be a $q \times n$ matrix, and let c_i be the ith row of C. The diagnostic matrix of $\{A, B, C, D\}$ will be arranged in the following order:

$$P = \begin{bmatrix} c_1 \\ c_1A \\ \vdots \\ c_1A^{n-1} \\ \vdots \\ c_q \\ c_qA \\ \vdots \\ c_qA^{n-1} \end{bmatrix}. \tag{1}$$

Now an algorithm will be introduced to find the first r linearly independent rows in P. This is achieved by a series of elementary transformations. Let

$$K_rK_{r-1} \cdots K_2K_1P \triangleq KP \tag{2}$$

where K_i are lower triangular matrices with all diagonal elements unity, and are obtained in the following manner. Let $p_1(j)$ be the first nonzero element from the left in the first row of P, where j denotes the position. Then K_1 is chosen so that all, except the first, elements of the jth column of K_1P are zero. Let $p_2(l)$ be the first nonzero element from the left of the second row of K_1P. Then K_2 is chosen so that all, except the first two, elements of the lth column of K_2K_1P are zero. Proceed in this manner until all linearly independent rows of P are found. Note that in this process, if one row is identically zero, then proceed to the next nonzero row. By multiplying these K_i, we obtain $K = K_rK_{r-1} \cdots K_2K_1$. Note that KP will be finally of form

Manuscript received February 21, 1973; revised May 21, 1973.
The author is with the Department of Electrical Sciences, State University of New York, Stony Brook, N.Y. 11790.